



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/607,370	06/30/2000	Reiner Kraft	ARC9-2000-0046-US1	8705

23334 7590 02/23/2004

FLEIT, KAIN, GIBBONS, GUTMAN, BONGINI
& BIANCO P.L.
ONE BOCA COMMERCE CENTER
551 NORTHWEST 77TH STREET, SUITE 111
BOCA RATON, FL 33487

EXAMINER

BURGE, LONDRA C

ART UNIT	PAPER NUMBER
----------	--------------

2178

DATE MAILED: 02/23/2004

3

Please find below and/or attached an Office communication concerning this application or proceeding.

dm

Office Action Summary

Application No.

09/607,370

Applicant(s)

KRAFT ET AL.

Examiner

Londra C Burge

Art Unit

2178

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 6/30/2000.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-20 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-20 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. §§ 119 and 120

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
* See the attached detailed Office action for a list of the certified copies not received.
- 13) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application) since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.
a) ☐ The translation of the foreign language provisional application has been received.
- 14) ☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121 since a specific reference was included in the first sentence of the specification or in an Application Data Sheet. 37 CFR 1.78.

Attachment(s)

- 1) ☒ Notice of References Cited (PTO-892) 4) ☐ Interview Summary (PTO-413) Paper No(s). _____
- 2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948) 5) ☐ Notice of Informal Patent Application (PTO-152)
- 3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ 6) ☐ Other: _____

DETAILED ACTION

1. This action is responsive to communications: original application filed 6/30/2000
2. Claims 1-20 are pending. Claims 1, 14 and 20 are independent claims.

Claim Rejections - 35 USC § 103

3. **The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:**

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. **Claims 1-3, 10, 14-16, and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), U.S. Patent No. 6,145,003 filed 12/17/1997 issued 11/7/2000, in view of Adar et al. (herein after Adar) U.S. Patent No. 6,493,702 B1 filed 5/5/1999 issued 12/10/2002.**

In regard to independent claim 1, Sanu teaches of a *method for browser-enhanced* (Sanu Col 1 Line 40) *web crawling* (Sanu Title) *associated with a network* (Sanu Abstract Line 2) *of hub processing units coupled to a plurality of information processing units over a network* (Sanu Col 1 line 11-13), *the method executed by a web crawler* (Sanu Abstract Line 8) *on a hub processing unit* (Sanu Col 3 Line 58) *associated with the network* (Sanu Abstract Line 2) *comprising the steps of: retrieving a document* (Sanu Col 5 Line 24-25) *at an address* (Sanu Col 5 Line 34); *loading secondary documents* (Sanu Col 10 Line 23); *sending to one or more information processing units* (Sanu Col 3 Lines 20-21) *a browser side script* (Sanu Col 7 Line 10) *to gather metadata* (Sanu Col 10 Line 11); *and performing the sub-steps of: producing a*

Art Unit: 2178

final HTML markup (Sanu Col 1 Line 33); *analyzing* (Sanu Col 5 Line 42) *and summarizing the final HTML markup* (Sanu Col 1 Line 33) *to produce metadata.* (Sanu Col 10 Line 11)

Sanu does not specifically mention the processing units being a hub or summarizing the final HTML markup. However, Adar teaches of a *Hub* (Adar Col 3 Line 20) and *summarizing* (Adar Col 13 Lines 7-8). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Adar to Sanu, providing Sanu the benefit of ensuring a hub processing unit, which is important for an exemplary operating environment as taught by Sanu Col 3-4. Also, providing Sanu the benefit of summarizing the final HTML before producing metadata about the document.

In regard to dependent claim 2, Sanu teaches *wherein the retrieving a document* (Sanu Col 5 Line 24-25) *at an address* (Sanu Col 5 Line 34) *step further comprises retrieving a document* (Sanu Col 5 Line 24-25) *at an address* (Sanu Col 5 Line 34) *selected from the group of addresses* (Sanu Col 5 Line 34) *consisting of a nodal address, a network address* (Sanu Col 2 Line 54), *a URL* (Sanu Col 6 Line 37) *and equivalents.*

In regard to dependent claim 3, Sanu teaches *wherein the analyzing* (Sanu Col 5 Line 42) *and summarizing step* as claimed in claim 1 *further comprises analyzing* (Sanu Col 5 Line 42) *and summarizing the whole and complete* (Sanu Col 8 Line 58 i.e. entire) *document.* (Sanu Col 9 Line 32)

Sanu does not specifically mention summarizing. However, Adar mentions *summarizing* (Adar Col 13 Lines 7-8). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Adar to Sanu, providing Sanu the benefit of summarizing the entire document to produce the metadata information.

In regard to dependent claim 10, Sanu teaches *wherein the tertiary sub-step of crawling* (Sanu Abstract) *further comprises the sub-steps of: issuing an HTTP command* (Sanu Col 6 Line 65) *to a web server* (Sanu Col 13 Line 23) *named* (Sanu Col 9 Line 3) *in the URL* (Sanu Col 9 Line 2); *receiving contents* (Sanu Col 3 Line 29) *of an HTML page* (Sanu Col 7 Line 8) *as a result* (Sanu Col 5 Line 43) *of the issued HTTP command* (Sanu Col 6 Line 65); *and passing* (Sanu Col 7 Line 39) *on the contents* (Sanu Col 5 Line 24) *of the HTML page* (Sanu Col 7 Line 8) *to a Page Rendering* (Sanu Col 7 Line 17) *subroutine*.

In regard to independent claim 14, claim 14 reflects the browser-enhanced web crawling as claimed in claim 1, and further view of the following, is rejected along the same rationale.

Sanu teaches *a computer readable medium* (Sanu Col 3 Line 34) *including programming instructions* (Sanu Col 3 Line 39).

In regard to dependent claim 15, claim 15 reflects retrieving a document at an address as claimed in claim 2, and in further view of the following, is rejected along the same rationale.

Sanu teaches *a computer readable medium* (Sanu Col 3 Line 34) and *instructions* (Sanu Col 3 Line 39).

In regard to dependent claim 16, claim 16 reflect the analyzing and summarizing as claimed in claim 3, and in further view of the following, is rejected along the same rationale.

Sanu teaches *a computer readable medium* (Sanu Col 3 Line 34) and *instructions* (Sanu Col 3 Line 39).

In regard to independent claim 20, Sanu teaches *a browser-enhanced* (Sanu Col 1 Line 40) *web crawling unit* (Sanu Title) *associated with a network* (Sanu Abstract Line 2) *of a*

Art Unit: 2178

plurality of hub processing units coupled to a plurality of information processing units over a network (Sanu Col 1 Line 11-13), the browser enhanced (Sanu Col 1 Line 40) web crawling (Sanu Title) unit on a hub processing unit comprising: a retrieval unit (Sanu Abstract 1-2) for retrieving a document (Sanu Col 5 Line 24-25) at an address (Sanu Col 5 Line 34); a loader for loading secondary documents as required (Sanu Col 10 Line 23), an output (Sanu Col 1 Line 64) for sending to one or more information processing units (Sanu Col 3 Line 20-21) a browser side script (Sanu Col 7 Line 10) to gather metadata (Sanu Col 10 Line 11); a producer for producing a final HTML markup (Sanu Col 1 Line 33); and a summarizer for analyzing and summarizing the final HTML markup (Sanu Col 1 Line 33) to produce the final metadata (Sanu Col 10 Line 11).

Sanu does not specifically mention a summarizing for analyzing and summarizing. However Adar mentions analyzing and summarizing. (Adar Col 13 Lines 7-8) It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Adar to Sanu, providing Sanu the benefit of summarizing that can be based on a concatenation of selected documents as taught by Adar Co 13 Lines 7-12

5. Claims 4-6, 17-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), as claimed in claim 1, in view of Adar et al. (herein after Adar) as claimed in claim 1, and in further view of Aganovic et al. (herein after Aganovic) U.S. Patent No. 6,105,042 filed 2/13/1998 issued 8/15/2000.

In regard to dependent claim 4, Sanu teaches *the step of analyzing* (Sanu Col 5 Line 42) *any image data present* (Sanu Col 7 Line 15) *in the document* (Sanu Col 7 Line 25) *and any*

Art Unit: 2178

image data present (Sanu Col 5 Line 42) in the documents (Sanu Col 7 Line 25) utilizing optical character recognition techniques.

Sanu does not specifically teach of optical character recognition. However, Aganovic teaches *optical character recognition* (Aganovic Col 7 Line 3-4). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Aganovic to Sanu, providing Sanu the benefit of applying optical character recognition to prepare data for document storage as taught by Aganovic Column 7 Line 1.

In regard to dependent claim 5, Sanu teaches *wherein the step of loading secondary documents (Sanu Col 10 Line 23) further comprises the loading of secondary documents (Sanu Col 10 Line 23) including documents selected from the group of documents consisting of in-line frames, frames, images, image maps, applets, audio, video or equivalents.*

Sanu does not specifically mention in-line frames, frames, images, image maps, applets, audio, video or equivalents. However, Aganovic teaches *in-line frames, frames* (Aganovic Col 8 Line 64), *images* (Aganovic Col 8 Line 49-50), *image maps* (Aganovic Col 1 Line 29 i.e. bitmap), *applets* (Aganovic Col 9 Line 65), *audio* (Aganovic Col 10 Line 31), *video* (Aganovic Col 10 Line 31) *or equivalents*. It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Aganovic to Sanu, providing Sanu the benefit of applying frames to enable the user to control viewing as taught by Aganovic Col 8 Line 66-67.

In regard to dependent claim 6, Sanu teaches *wherein the step of analyzing (Sanu Col 5 Line 42) any image data present (Sanu Col 7 Line 15) in the document (Sanu Abstract Line 2) and any image data present (Sanu Col 7 Line 15) in the documents (Sanu Abstract Line 2) utilizing optical character recognition techniques further comprises analyzing (Sanu Col 5 Line*

Art Unit: 2178

42) *any images and image maps in the image data* (Sanu Col 7 Line 15) *to product text data.*
(Sanu Col 7 Line 7)

Sanu does not specifically mention optical character recognition techniques, images and image maps. However, Aganovic teaches *optical character recognition* (Aganovic Col 7 Lines 3-4) *techniques, images* (Aganovic Col 8 Lines 49-50) *and image maps.* (Aganovic Col 1 Line 29 i.e. bitmap) It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Aganovic to Sanu, providing Sanu the benefit of applying optical character recognition to prepare data for document storage as taught by Aganovic Column 7 Line 1, and the benefit of apply images and image maps to determine the type pf presentations document desired by the user as taught by Aganovic Col 8 Lines 49-50.

In regard to dependent claim 17, claim 17 reflects the image analyzing as claimed in claim 4, and I further view of the following, is rejected along the same rationale.

Sanu teaches *a computer readable medium* (Sanu Col 3 Line 34) and *instructions* (Sanu Col 3 Line 39).

In regard to dependent claim 18, claim 18 reflects the loading of secondary instructions as claimed in claim 5, and in further view of the following, is rejected along the same rationale.

Sanu teaches *a computer readable medium* (Sanu Col 3 Line 34) and *instructions* (Sanu Col 3 Line 39).

In regard to dependent claim 19, Sanu teaches *wherein the analyzing* (Sanu Col 5 Line 42) *instructions* (Sanu Col 3 Line 39) *for analyzing* (Sanu Col 5 Line 42) *any image data present* (Sanu Col 7 Line 15) *in the document* (Sanu Abstract Line 2) *and any image date present in the documents* (Sanu Abstract 3 Line 2) *utilizing optimal character recognition techniques further*

comprises analyzing (Sanu Col 5 Line 42) instructions (Sanu Col 3 Line 39) for analyzing (Sanu Col 5 Line 42) any images and image maps in the image data (Sanu Col 7 Line 15) to produce text data. (Sanu Col 7 Line 7)

Sanu does not specifically teach of images, image maps and utilizing optimal character recognition techniques. However, Aganovic teaches of *utilizing optimal character recognition techniques* (Aganovic Col 7 Line 3-4), *images* (Aganovic Col 8 Line 49-50), and *image maps* (Aganovic Col 1 Line 20 i.e. bitmap) It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Aganovic to Sanu, providing Sanu the benefit of applying optical character recognition to prepare data for document storage as taught by Aganovic Column 7 Line 1, and the benefit of apply images and image maps to determine the type of presentations document desired by the user as taught by Aganovic Col 8 Lines 49-50.

6. **Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), as claimed in claim 1, in view of Adar et al. (herein after Adar) as claimed in claim 1, in further view of Meyerzon et al. (herein after Meyerzon 6,199,081) U.S. Patent No. 6,199,081 B1 filed 6/30/1998 issued 3/6/2001, and in further view of Meyerzon et al. (herein after Meyerzon 6,638,314) U.S. Patent No. 6,638,314 B1 filed 6/26/1998 issued 10/28/2003.**

In regard to dependent claim 7, Sanu teaches *wherein the retrieving step* (Sanu Col 5 Line 24-25) *further comprises performing the sub-steps of: initializing a first list* (Sanu Col 6 Line 33) *with seed values* (Sanu Col 6 Line 34); *checking* (Sanu Col 7 Line 40) *if there are any URLs to be processed* (Sanu Col 7 Line 44) *and if there are, performing the secondary sub-steps*

Art Unit: 2178

of: determining (Sanu Col 9 Line 45) if a URL is in a second list (Sanu Col 5 Line 67 and Col 6 Lines 1-3); and if it is not in the second list (Sanu Col 5 Line 67 and Col 6 Lines 1-3); then performing the tertiary sub-steps of: inserting the URL into the first list (Sanu Col 6 Line 45-46); scheduling the URL for crawling; crawling the URL when scheduled to do so; removing the URL from the first list (Sanu Col 5 Line 67 and Col 6 Lines 1-3) after the scheduled crawling; entering the URL into the second list (Sanu Col 5 Line 67 and Col 6 Lines 1-3); and repeating (Sanu Col 10 Line 7) the checking step (Sanu Col 2 Line 22) until there are no more URLs to be processed; where if the determining step (Sanu Col 9 Line 45) determines that the URL is in the second list (Sanu Col 5 Line 67 and Col 6 Lines 1-3) then repeating (Sanu Col 10 Line 7) the checking step (Sanu Col 7 Line 40) until there are no more URLs to be processed.

Sanu does not specifically teach of scheduling the URL for crawling, crawling the URL when scheduled to do so and scheduled crawling. However Meyerzon 6,199,081 B1 teaches *scheduling the URL for crawling* (Meyerzon 6,199,081 B1 Col 9 Lines 10-11) and *crawling the URL when scheduled to do so* (Meyerzon 6,199,081 B1 Col 9 Lines 20-21) and *scheduled crawling* (Meyerzon 6,199,081 B1 Col 9 Lines 10-11). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon 6,199,081 B1 to Sanu, providing Sanu the benefit of applying starting URLs that instruct the gather process where to begin as taught by Meyerzon 6,199,081 B1 Col 9 Lines 10-11.

Sanu does not specifically teach of there being no more URLs to be processed. However Meyerzon 6,638,314 B1 teaches of *no more URLs to be processed* (Meyerzon 6,638,314 B1 Col 12 Line 57). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon 6,638,314 B1 to Sanu, providing Sanu the benefit of

Art Unit: 2178

processing linked URLs within the filtered data corresponding to an electronic document to be complete as taught by Meyerzon 6,638,314 B1 Col 13 Line 58-59

7. **Claims 8-9 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), as claimed in claim 1, in view of Adar et al. (herein after Adar) as claimed in claim 1, as in further view of Meyerzon et al. (herein after Meyerzon) U.S. Patent No. 6,199,081 B1 filed 6/30/1998 issued 3/6/2001.**

In regard to dependent claim 8, Sanu teaches *wherein the sub-step of initializing a first list (Sanu Col 6 Line 33) with seed values (Sanu Col 6 Line 34) further includes the list being a URL pool*

Sanu does not specifically teach of a URL pool. However, Meyerzon teaches of a *URL Pool* (Meyerzon Col 9 Line 34 i.e. list of all URLs). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of applying a URL Pool, which maintains a list of URLs that are currently being processed or have not yet been processed, as taught by Meyerzon Column 9 Lines 22-24.

In regard to dependent claim 9, Sanu teaches *wherein the sub-step of determining if a URL is in a second list (Sanu Col 5 Line 64 and Col 6 Line 2-3) further includes the second list being (Sanu Col 5 Line 64 and Col 6 Line 2-3) a visited pool.*

Sanu does not specifically teach of a visited pool. However, Meyerzon teaches a *visited pool* (Meyerzon Col 9 Lines 34-36 i.e. list of all URLs that have been visited). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of applying a visited pool that contain a list of all

Art Unit: 2178

URLs that have been visited or attempted to be visited during either the current Web crawl or a previous Web crawl.

8. **Claims 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), as claimed in claims 1 and 10, in view of Adar et al. (herein after Adar) as claimed in claim 1, as in further view of Meyerzon et al. (herein after Meyerzon) U.S. Patent No. 6,638,314 B1 filed 6/26/1998 issued 10/28/2003.**

In regard to dependent claim 11, Sanu teaches further including the sub-steps performed by the Page Rendering (Sanu Col 7 Line 17) subroutine comprising: receiving the contents (Sanu Col 3 Line 19) of the HTML page (Sanu Col 7 Line 8) in the Page Rendering (Sanu Col 3 Line 39) subroutine; building an in-memory (Sanu Col 6 Line 52) representation of a Layout for the HTML page (Sanu Col 7 Line 8) and if more data is needed to properly form the representation, then performing the sub-steps of: requesting additional web-based information (Sanu Abstract Line 1-2); gathering this additional web-based information (Sanu Abstract Line 1-2); inserting any URLs associated with this additional web-based information (Sanu Abstract Line 1-2) into the second list (Sanu Col 5 Line 67 and Col 6 Line 2-3) and a URL (Sanu Col 9 Line 2) cache (Sanu Col 6 Line 52); building (Sanu Col 9 Line 29) a final amended representation; and forwarding the final amended representation to an Extraction subroutine; wherein, if no more data is needed to properly form the in-memory (Sanu Col 6 Line 52) representation, then forwarding the in-memory (Sanu Col 6 Line 52) representation to the Extraction subroutine.

Sanu does not specifically mention representation of HTML page, inserting URLs, final amended representation, forwarding and extraction subroutine. However, Meyerzon teaches *representation of a Layout* (Meyerzon Col 5 Lines 35-36), *amended representation* (Meyerzon Col 5 Lines 35-36 i.e. changed), *requesting* (Meyerzon Col 22 Lines 52), *gathering* (Meyerzon Col 11 Lines 49), *inserting any URLs* (Meyerzon Col 8 Lines 67), *forwarding* (Meyerzon Col 16 Lines 59 i.e. sending), *Extraction subroutine* (Meyerzon Col 5 Lines 8), and *no more data* (Meyerzon Col 13 Lines 57). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of gathering and requesting information to be used and inserting any URLs associated with the information and amended representation i.e. changes to documents which would includes an existing index and history map as its transaction log contained in the history map as taught by Meyerzon Col 5 Lines 35-42.

9. **Claims 12 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sanu et al. (herein after Sanu), as claimed in claims 1 and 10, in view of Adar et al. (herein after Adar) as claimed in claim 1, Meyerzon et al. (herein after Meyerzon 6,199,081) as claimed in claim 7, in view of Meyerzon et al. (herein after Meyerzon 6,638,314 B1) as claimed in claim 11, in further view of Hughes et al (herein after Hughes) U.S. Patent No. 5,892,908 filed 9/10/1996 issued 4/6/1999 and in further view of Aganovic et al. (herein after Aganovic) U.S. Patent No. 6,105,042 filed 2/13/1998 issued 8/15/2000.**

In regard to dependent claim 12, Sanu teaches *further including the sub-steps performed by the Page Extraction subroutine comprising: accessing a set of memory structures*

Art Unit: 2178

(Sanu Col 5 Line 4) *of the Page Renderer* (Sanu Col 7 Line 17); *copying a text portion* (Sanu Col 6 Line 52) *of the structures* (Sanu Col 5 Line 4) *into a text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42); *inspecting any in-line GIF and JPEG image references* (Sanu Col 5 Line 31) *in the memory structures* (Sanu Col 5 Line 4); *extracting alternate text attributes; adding the alternate text attributes to a text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42); *invoking an optical character recognition engine; analyzing any in-line GIF and JPEG images using the optical character recognition engine for text* (Sanu Col 4 Line 47) *content* (Sanu Col 5 Line 43); *extracting text* (Sanu Col 4 Line 47) *content* (Sanu Col 5 Line 43) *from the GIF and JPEG images; adding text* (Sanu Col 4 Line 47) *content* (Sanu Col 5 Line 43) *from the images* (Sanu Col 7 Line 15) *to the text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42); *and forwarding the text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42) *to a Page Summarizer subroutine.*

Sanu does not specifically mention a Page Extraction subroutine, copying, extracting alternate text attributes, adding the alternate text attributes, extracting, and forwarding. However, Meyerzon 6,638,314 B1 mentions *Page Extraction subroutine* (Meyerzon 6,638,314 B1 Col 5 Line 8), *copying* (Meyerzon 6,638,314 B1 Col 15 Line 15), *extracting* (Meyerzon 6,638,314 B1 Col 5 Line 8) *alternate* (Meyerzon 6,638,314 B1 Col 4 Line 11 i.e. others) *text attributes* (Meyerzon 6,638,314 B1 Col 16 Line 11), *adding* (Meyerzon 6,638,314 B1 Col 4 Line 3) *the alternate* (Meyerzon 6,638,314 B1 Col 4 Line 11 i.e. others) *text attributes* (Meyerzon 6,638,314 B1 Col 16 Line 11), *extracting* (Meyerzon 6,638,314 B1 Col 5 Line 8), *and forwarding* (Meyerzon 6,638,314 B1 Col 16 Line 56 i.e. sending). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of extracting data to store in an index taught by Meyerzon 6,638,314 B1 Col 5

Art Unit: 2178

Line 8, copying history map entries into a transaction log taught by Meyerzon 6,638,314 B1 Col 15 Line, text attributes to retrieve electronic document taught by Meyerzon 6,638,314 B1 Col 16 Line 14.

Sanu does not specifically teach of images, image maps and utilizing optimal character recognition techniques. However, Aganovic teaches of *utilizing optimal character recognition techniques* (Aganovic Col 7 Line 3-4). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Aganovic to Sanu, providing Sanu the benefit of applying optical character recognition to prepare data for document storage as taught by Aganovic Column 7 Line 1, and the benefit of apply images and image maps to determine the type of presentations document desired by the user as taught by Aganovic Col 8 Lines 49-50.

Sanu does not specifically mention JPEG and GIF images. However, Hughes mentions *JPEGs and GIF* (Hughes Col 2 Lines 49). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Hughes to Sanu, providing Hughes the benefit of having JPEGs and GIF for files that can be in a number of formats as taught by Hughes Col 2 Line 47.

Sanu does not specifically mention Page summarizing. However, Adar mentions *summarizing* (Adar Col 13 Lines 7-8). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Adar to Sanu, providing Sanu the benefit of a document summary generated based on a concatenation of selected documents as taught by Adar Col 13 Line 8-11.

In regard to dependent claim 13, Sanu teaches *the Page Summarizer subroutine comprising: receiving* (Sanu Col 5 Line 24-25) *a text* (Sanu Col 7 Line 47) *map* (Sanu Col 7

Art Unit: 2178

Line 42) *from the Page Extractor subroutine; processing* (Sanu Col 1 Line 54-55) *the text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42) *in an application-specific manner* (Sanu Col 4 Line 26-27); *applying data extraction patterns to the text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42); *translating resultant data from the applying step* (Sanu Col 9 Line 16 i.e. applied); *forwarding any URLs present in the text* (Sanu Col 7 Line 47) *map* (Sanu Col 7 Line 42) *to a manager subroutine; and forwarding any extracted data and metadata* (Sanu Col 10 Line 11) *to application logic* (Sanu Col 4 Line 26-27).

Sanu does not specifically mention Page summarizing. However, Adar mentions *summarizing* (Adar Col 13 Lines 7-8). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Adar to Sanu, providing Sanu the benefit of a document summary generated based on a concatenation of selected documents as taught by Adar Col 13 Line 8-11.

Sanu does not specifically mention a Page Extraction subroutine, copying, extracting alternate text attributes, adding the alternate text attributes, extracting, and forwarding. However, Meyerzon 6,638,314 B1 mentions *Page Extraction subroutine* (Meyerzon 6,638,314 B1 Col 5 Line 8). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of extracting data to store in an index taught by Meyerzon 6,638,314 B1 Col 5 Line 8.

Sanu does not specifically mention a manager subroutine. However Meyerzon 6,638,314 B1 mentions a management routine (Meyerzon 6,638,314 B1 Col 12 Line 42 i.e. control). It would have been obvious to one of ordinary skill in the art at the time the invention was made to

Art Unit: 2178

apply Meyerzon to Sanu, providing Sanu the benefit of adding a control step where the document or URL is retrieved as taught by Meyerzon 6,638,314 B1 Col 12 Line 35-53.

Sanu does not specifically mention translating resulting data. However, Meyerzon 6,199,081 B1 mentions translations (Meyerzon 6,199,081 B1 Col 17 Lines 15 i.e. interpret) resulting data (Meyerzon 6,199,081 B1 Col 17 Line 40). It would have been obvious to one of ordinary skill in the art at the time the invention was made to apply Meyerzon to Sanu, providing Sanu the benefit of interpreting and message so proper action can be taken as taught by Meyerzon 6,199,081 B1 Col 17 Line 16.

Conclusion

10. **The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.**

Acton et al.	US Patent No. 6,094,532	issued	7/25/2000
Agrawal et al.	US Patent No. 6,233,575 B1	issued	5/15/2001

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Londra C Burge whose telephone number is 703-305-8784. The examiner can normally be reached on 8:30am to 5:00pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather Herndon can be reached on 703-308-5186. The fax phone number for the organization where this application or proceeding is assigned is 703-746-7239.

Art Unit: 2178

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is 703-305-3900.

Any response to this action should be mailed to:

Commissioner of Patents and Trademarks
Washington, DC 20231

Or faxed to:

(703) 746-7239 (for formal communications intended for entry)

Or:


(703) 746-7240 (for informal or draft communications, please label
"PROPOSED" or "DRAFT")

Or:

(703) 746-7238 (for after-final communications)

**Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive,
Arlington, VA, Fourth Floor (Receptionist).**

Londra Burge
2/12/2004


STEPHEN S. HONG
PRIMARY EXAMINER